Geological Survey of Finland
MTR
Mineral Economy Solutions
Geological Survey of Finland (GTK)
PO Box 96, Vuorimiehentie 5
FI-02151 Espoo
FINLAND

December 26, 2022 9/2023

Estimating grade and tonnage models of mineral deposit data using statistical imputation: an improved workflow for undiscovered mineral resource assessments

Leonardo Feltrin and Martina Bertelli

GTK Open File Work Report

GEOLOGICAL SURVEY OF FINLAND **DOCUMENTATION PAGE**

26.12.2022

Authors Leonardo Feltrin Martina Bertelli	Type of report Open File Work Report Commission by
Title of report Estimating grade and tonnage models of mineral deposit data using statistical imputation: an improved workflow for undiscovered mineral resource assessments	

Abstract: Missing data is a frequent issue in geoscience. Mineral resource assessment data are prone to this problem with incomplete information occurring in records of tonnage and grade of mineral deposits. We review several data imputation methodologies, the classification of missing information (MI) types, and define a workflow that evaluates usage of imputation in quantitative undiscovered mineral resources assessments (UMRA) when constructing grade and tonnage curves. UMRA is now targeted at technology metals and other industrial minerals of the future economies; thus, it has implication for sustainability and carbon neutrality. Results obtained on Au mineralization data (Chugach-type lode gold veins) in Alaska, illustrate how modern imputation requires simulations to evaluate the most appropriate imputation algorithm, given the bias introduced in the regression model by the specific input dataset. More than ~118,800 simulations were carried out to evaluate the behavior of five imputation methodologies and their comparison against the results of listwise deletion. The results indicate that imputation performance and algorithm convergence are controlled by the distribution, amount, and type of missing information, while also being partly a function of the simulation parameters. The simulations confirm the limitations of standard linear regression suggesting that multiple imputation is a more appropriate methodology, especially when missing information is around 35% of the total data rows—which is undesirable, increasing the uncertainty of the estimates.

Keywords Imputation, Missing Information, Grade and Tonnage Models, Economics			
Geographical area Chugach National Forest, Alaska, USA (Data Source, USGS)			
Other information Research conducted as part of the internally funded Bedrock and Mineral Intelligence project (WP2)			
Report serial Open File Work Report		Archive code 9/2023	
Total pages 16	Language English	Price	Confidentiality Public
Unit MTR5		Project code 50402-2010622	
Signature/name Leonardo Feltrin		Signature/name Tero Niiranen	

Contents

1 INTRODUCTION

Imputation of data is an important field of statistical analysis and is represented by a process of “filling the data gaps” where data are either, missing, incorrect or inconsistent because of errors in their editing or unwanted external bias. Imputation comprises both the determination of missing values and the replacement value mechanism—defining what needs to be imputed in the gaps. The two are fundamental to ensure that estimates are of high quality and plausible (Ferguson and Winkler 2000).

If we consider the statistician methodological perspective (Cheema 2014; Enders 2010) the problem of missing information (MI) has been subject of debate and development for nearly a century (e.g., Wilks already in 1932 proposes the earliest application of maximum likelihood imputation in bivariate analysis). Despite the long history, major advances and statistical tools were developed in the 1970s in response to the digital revolution. These later developments considered more complex imputation methods, which marked sharp improvements in the methodology. During this period, Rubin (1976) defined a theoretical framework for MI problems that is still valid today (e.g., Madani and Bazarbekov 2021). If we survey the range of imputation techniques (e.g., Enders 2010), we can find a multitude of approaches. In this context, a division exists between techniques that work with data characterized by randomly distributed MI and techniques that work with data characterized by systematic distributions of MI, controlled by variable(s)-missingness interdependences. This latter case is common in mineral exploration and mining datasets. Often the data are not missing completely at random (MCAR), but rather missing not at random (MNAR); see Rubin (1976) for appropriate definitions and da Silva and Costa (2019) for an application.

Choosing an appropriate method is partly directed by the distribution of MI, and in recent time it has been restricted to the most advanced methods of maximum likelihood (Enders 2010) and multiple imputation (van Buuren and Groothuis-Oudshoorn 2011), with the limited exception of using stochastic regression imputation algorithms (Enders 2010; van Buuren and Groothuis-Oudshoorn 2011) which however do not always perform well with missing at random (MAR) and MNAR data (Enders 2010). Once the listed methodologies are implemented, they provide a reconstruction of the data distribution that partly solves the problem of missingness (in the sense that the replaced values might be still different from the true, real values), often resulting in a bias that can be measured if multiple simulations involving sets of amputations and subsequent imputations are carried out to explore solution variances against the original data.

Using an approach comparable to earlier work conducted on the assessment of imputation methodologies, synthetic data were generated with variable parametrizations (e.g., different types of MI) to simulate different missingness scenarios and allow sensitivity analysis to understand the effect of missingness and its replacement with imputed values, when considering geological data distributions representing grade and tonnage of ore deposits. Results were compared against a regression model developed on the mineral resources dataset (represented by complete-case data), to obtain an appreciation of the quality of imputations performed. The final objective of minimizing bias together with the need of identifying appropriate imputation algorithms were key aspects to this analysis. All these tests are useful to characterize the behavior of imputation methods and assist the selection of an appropriate imputed dataset, subsequently used to construct a grade and tonnage model, an important phase of the three-part method (Singer 1993; Singer and Menzie 2010) in undiscovered mineral resources assessments (UMRA).

This report illustrates an experimental workflow that attempts the integration of imputation with the representation of grade and tonnage data in logistic form. A summary of the simulations conducted on synthetic data is proposed. Simulations were generated from original geological grade and tonnage information derived from a public dataset on Au deposits sited in the Chugach region, located in central-southern Alaska, USA (Figures 1, 2).

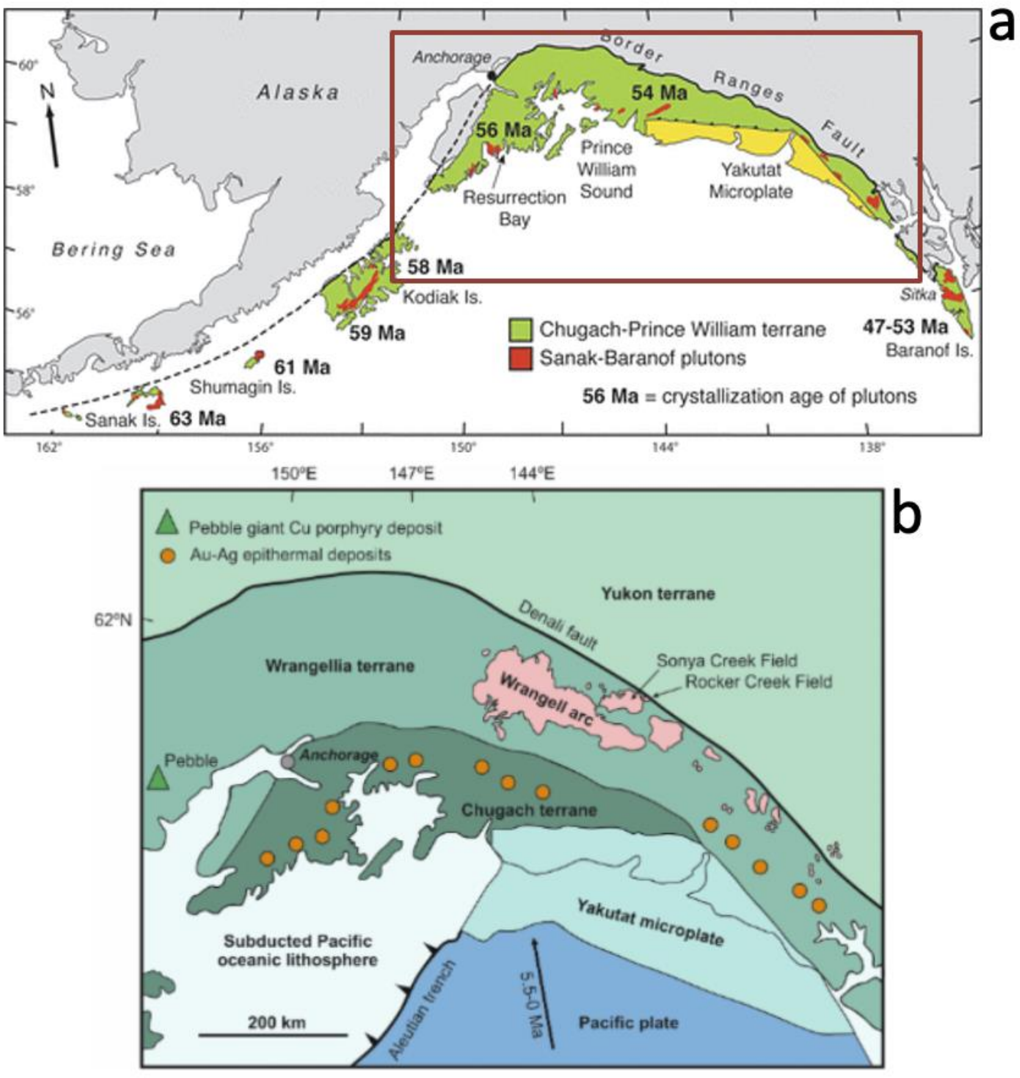

Figure 1: Tectonic plutonism on the Chugach-Prince William terrane (a), and geological representation of the plate margins and overall arc-contractional setting of the Chugach and adjacent accretionary terranes. Porphyry and epithermal Au-Ag mineralization (b) occur in both the Chugach and Wrangellia terranes, adapted from (Kepezhinskas et al. 2022).

Further details of these experiments are explained below in the methodology and simulations sections. In the final part, the report presents some of the most relevant results obtained by constructing multiple grade fitting models for a univariate missing information case that most closely resemble the original data missingness with only Ag grades being affected by MI. More in depth analysis and discussion is then provided to explain the implications of these experiments when considering grade and tonnage data and their statistical estimation.

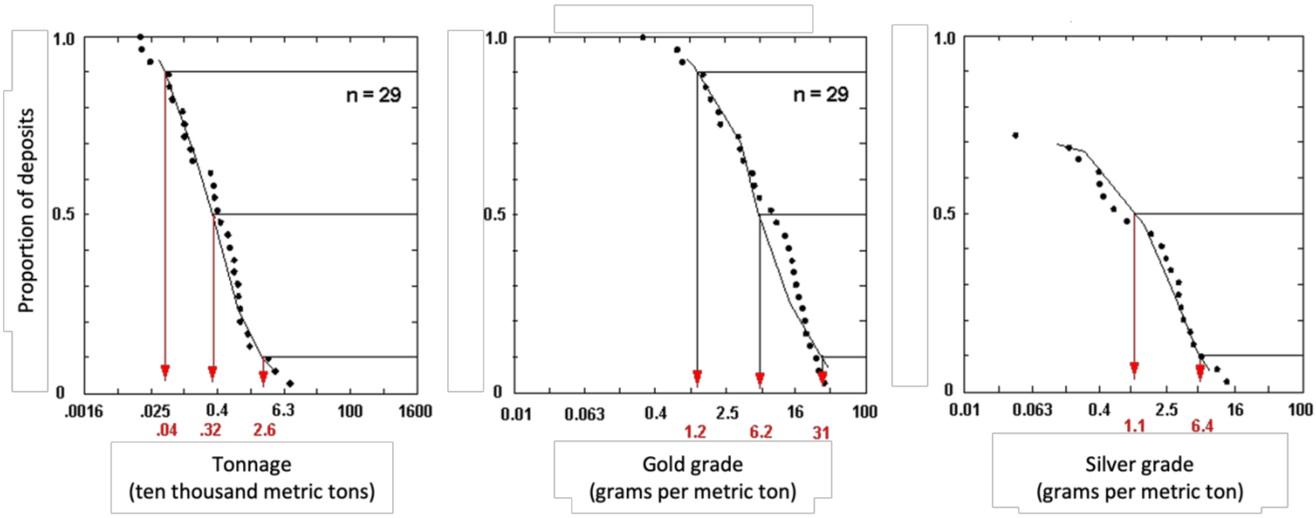

Figure 2: Tonnages and grades of Chugach-type low- sulfide Au-quartz vein deposits, after Bliss (2004).

2 METHODOLOGY

The workflow illustrated in Figure 3 includes two distinct phases. Firstly, a set of simulations carried out on multiple synthetic datasets employing the packages MICE (van Buuren and Groothuis-Oudshoorn 2011) and "norm2" (Schafer 2021; Takahashi 2017) to perform different imputation routines. This phase helps with algorithm's selection and evaluation. The second phase instead focuses on the construction of grade and tonnage curves based on the imputed data and their comparison against the complete-cases model and the listwise deletion model. In section 2.2, we discuss some of the differences and applicability. Finally, simulations of MI imputation algorithms were compared to listwise deletion (using the omit.na R-package of Figure 4 to remove MI rows).

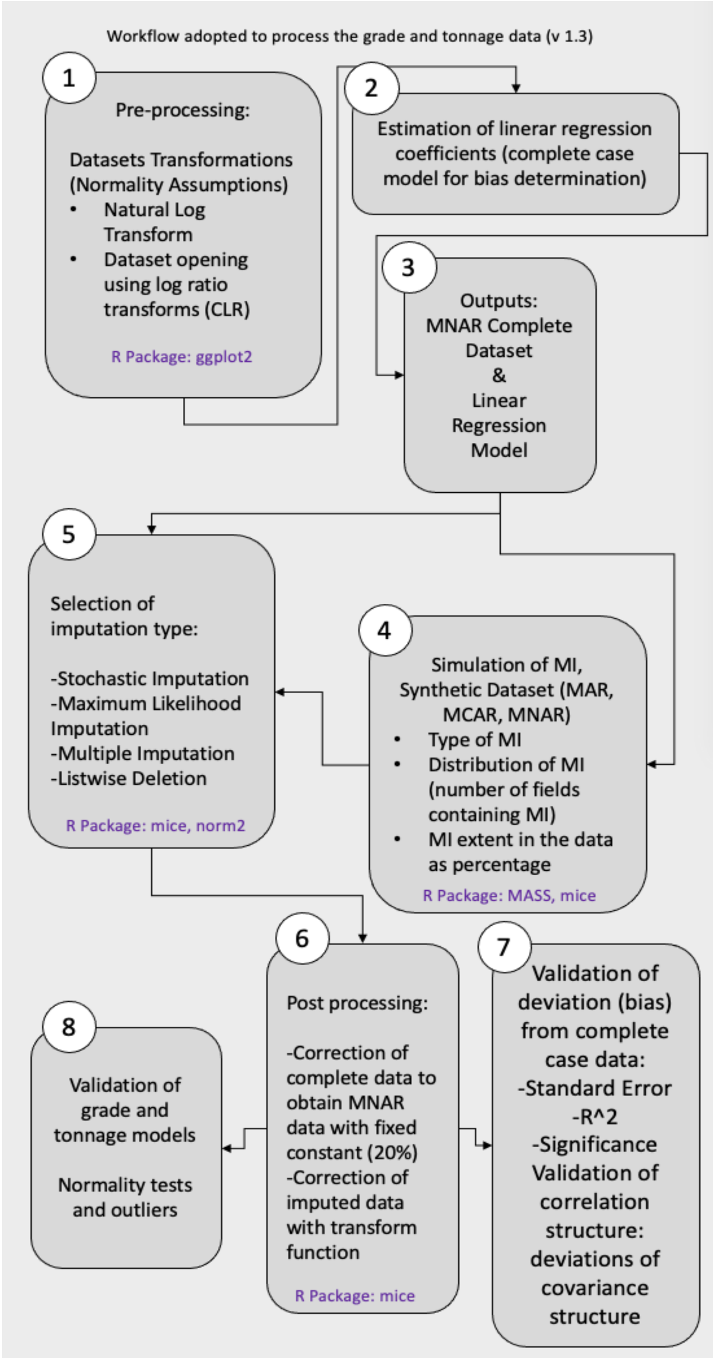

Figure 3: Proposed workflow adopted to merge imputation statistics with UMRA. Phases 1 to 5 represent simulation stages. From 6 to 8 the workflow considers the validation of the results against listwise deletion and integration of selected imputation results within grade and tonnage estimation curves.

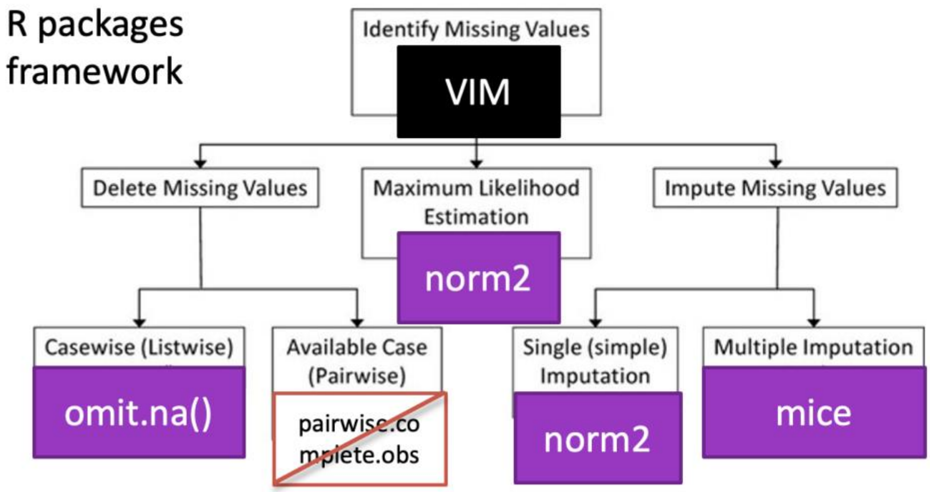

Figure 4: Schematic of R packages implemented in the simulations with specific imputation methods and to document the distributions of missing information and relative value replacements via imputation or row omissions when considering the listwise deletion approach.

2.1 Classification of missing information

Three district mechanisms of MI are commonly recognized (Rubin 1976; Enders 2010). If the MI is completely independent (a random distribution of gaps) it is commonly considered a missing completely at random (MCAR) case. A geological example could consider several geologists that worked on a drilling database and did not completely record the logging details (Geekiyanage et al. 2020). Another more structured distribution is the missing at random (MAR) case. This is a case of dependency of MI in a variable on a single or multiple external, measured variables, but not on the variable itself (e.g., Enders 2010). The geologists may decide to send to the lab for chemical analysis specific intervals of core, and this will be commonly decided upon examination of specific logging information (e.g., available rock classifications). Finally, missing not at random (MNAR) considers the case of MI on a variable as dependent on the variable itself (e.g., Enders, 2010). For instance, a geologist may decide to assay an interval of core that is close to existing mineralization. Recorded ore grades are then influencing the distribution of MI and the planning and design of geochemical studies (da Silva and Costa 2019).

MAR, MCAR and MNAR distributions were reconstructed using the *produce_NA* R package, see Figure 5 for an explanation of the different types of MI. A total of 2 cycles with 100 and 1000 simulations for each parametrization were generated repeatedly over synthetic MI data to evaluate the improvement of algorithm convergence and relative bias minimization (this report discusses only the first cycle of analyses with 100 simulation rounds). Simulation rounds considered the effect of MI as well as different sampling rates (Sim. N = 25, 50, 100) of the posterior distribution (van Buuren and Groothuis-Oudshoorn 2011). Results of the first batch of 100 cycles is reported in Table 1. Before discussing the results, we introduce the 3 algorithms considered in the reported experiments and discuss a series of validation measures.

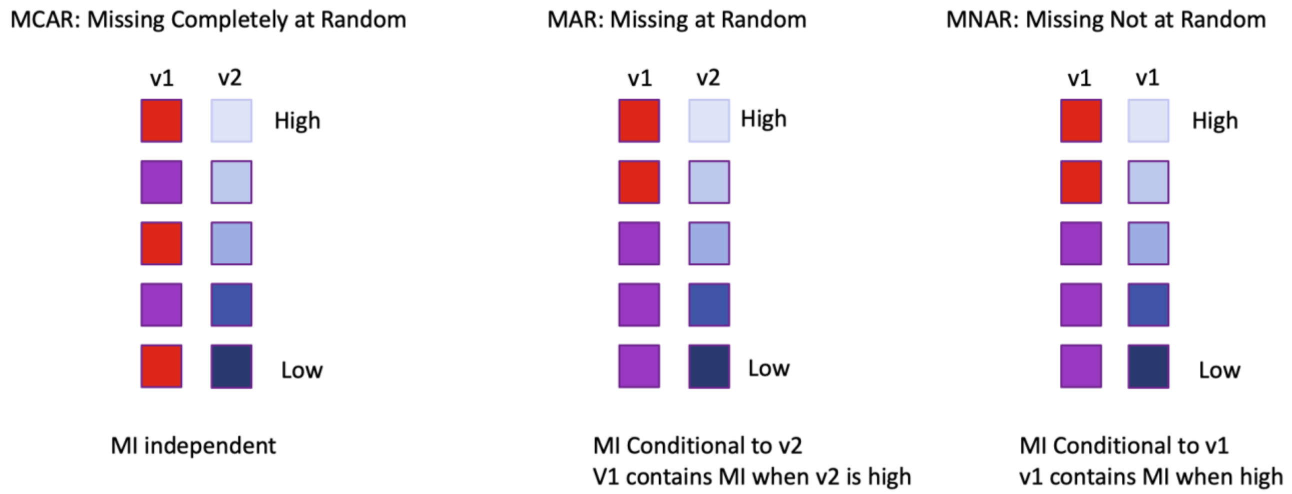

Figure 5: Two variable (v1, v2) examples of possible relationships of MI in variable v1 with a second variable (v2): MCAR case shows MI unrelated to v2 magnitude (high-low); MAR case shows v1 conditioned by v2 and constrained to high values of v2; MNAR case illustrates how MI in v1 is conditioned by the values of v1.

2.2 Imputation algorithms and validation

The MICE package (van Buuren and Groothuis-Oudshoorn 2011) offers a variety of imputation algorithms and a set of statistical measures of the quality of imputation performed using biases. Two univariate, linear regression algorithms were used in these simulations (norm.predict and norm.nob). They were compared with another method, predictive mean matching (pmm), that performs better when the correlation among variables is weak or non-significant (Morris et al. 2014), which often occurs when we have higher frequencies of MI gaps or if the MI distributions are cross-variable independent. The pmm algorithm employs multiple imputation with chained equations and is less dependent on the normality assumption, a condition typical of linear regression models. Multiple imputation (pmm) is expected to perform better than single pass imputation methods when the distribution of MI is of MAR or MNAR type (Enders 2010). These methodologies were also compared with parametric estimation models based on norm2 R package, which represents an example of maximum likelihood (ml) imputation (Schafer 1999, 2021).

Different statistical measures of bias and residual variance were used to allow models' comparison. Measures include *raw bias (RB), percent bias (PB), coverage rate* of the confidence interval *(CR), average width (AW) and root mean squared error (RMSE)*. The first two metrics express the difference (bias) between the expected values of the regression coefficients estimates, and what is considered the true value. *PB* rescales the result as a normalized percentage difference (see Table 1) with an upper limit of 5% (0-5% range results are considered acceptable). CR represents the proportion of confidence intervals that contain the true value. It should be equal or higher than the nominal value (95% in Table 1 or 0.95). AW indicates the statistical efficiency. It should be as narrow as possible however maintaining *CR* over the nominal value. *RMSE* represents a compromise between bias and variance; although widely in use, it was not used to compare the algorithms (refer to van Buuren 2018 for further details on RMSE limitations in imputation studies).

An additional test was carried out to evaluate a MNAR specific regression methodology, accounting for the common problem of under sampling and under representation of barren domains or low-grade poorly endowed areas (see Figure 6 for some examples of sampling bias occurring in mining and mineral exploration environments). These examples show that in most cases input data will be overestimations of the real crustal endowments of a commodity. It is customary to apply a correction to this deficiency by either empirical fixed rescaling of the data or alternatively as implemented in this analysis by applying a correction that relies on a linear fit (mnar.norm R package).

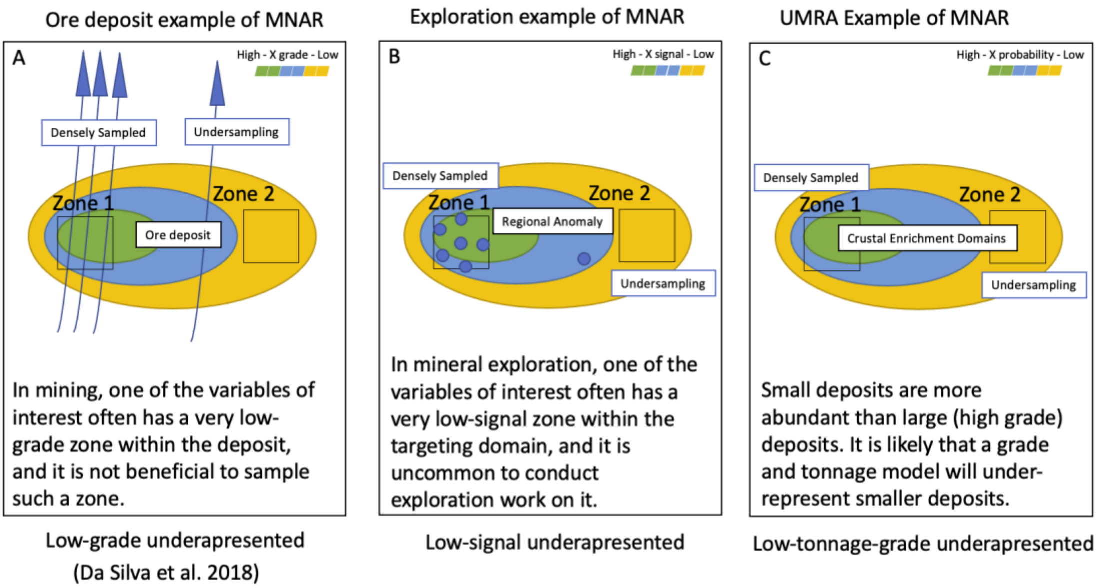

Figure 6. In many economic geology exploration and mining cases the distribution of missing data is conditioned by knowledge acquired during past exploration programs and mining activities. Conceptually when conducting resource definition studies (A) there is a tendency to oversample domains with proven resources. If we consider an exploration program (B) usually many surveys are carried out in areas that have shown potential representing a form of localization of exploration knowledge that biases missingness. The last example is an UMRA study (C), in most cases these assessments will report information on highly endowed terranes discounting areas that are assumed to be poorly endowed.

3 SIMULATIONS

The simulation performed considered a complex set of parameters presented in Table 1. We briefly outline the main findings based on the bias and variance measures discussed in 2.2 and observed in a variety of experiments.

3.1 Experimental setup

The statistical workflow included a series of pre-processing steps before simulation (Figure 3). The data were normalized using a log-transformation to approximate normality (a requirement and limitation of some of the imputation algorithms implementing linear regression or parametric modelling).

Complete-case data were used to develop a multiple-regression model that represented the true benchmark for the simulations. Biases and variances were then calculated by comparing the regression's fits done on multiple imputations outputs after amputation of the complete dataset (van Buuren 2018). A total of ~118,800 simulations, obtained from 216 parametrizations were run to evaluate sampling uncertainty of the posterior distribution and the variable distribution of MI in the synthetic data (Table 1 presents a preliminary summary of top results based on 100 simulations rounds).

3.2 Simulation results

The results obtained show a broad variability of biases and variances. These differences allowed the classification of cases that were likely representative of "good" imputations. If we restrict the analysis of the results to the summary in Table 1, we observe a range in RBs of one to two orders of magnitude around the zero-threshold value, indicating minimal bias and an inverse relationship with the sampling number *n* (larger sub-samples offer a better representation of the posterior distribution reducing bias of approximately one order of magnitude cfr. Rows 1 and 2 of Table 1 in the RB field or ~10% of the PBs). Coverage rates oscillate between a minimum of 0.26 and 0.97 with a total of 9 cases with either equal or higher CRs than the nominal value. All these results are either obtained with pmm or norm.nob imputations.

The largest AW bounds are found in the standard regression imputations (last three rows of Table 1) indicating complete failure of these solutions (close to infinite or unrealistically large bounds). The narrower AW are instead around 1.75 down to a minimum of 0.79 among the top 9 cases, with optimal PBs ~5% or smaller. AW were useful to assess the statistical efficiencies and showed that pmm is the most appropriate when missing information is distributed across all variables (see Figure 7 illustrating various examples of synthetic MI distributions across the considered variables v1-v3 representing incomplete grade and tonnage data). Stochastic regression imputation is however the best methodology for cases with a single variable affected by missingness. If we consider the type of MI (as expected) lower biases will occur in the most complete data with either MAR or MCAR type missingness.

Another set of simulations considered the improvement of modelling results in case of MNAR type data. Recent research indicated that if appropriate corrections are applied in the final part of the workflow this could provide a closer fit to the original complete case data (e.g., da Silva and Costa, 2019). These simulations included an assessment of bias (Figures 8 and 9) to evaluate the variation in model fit, depending on the algorithm of choice. Five distinct methodologies were used with MNAR type simulations.

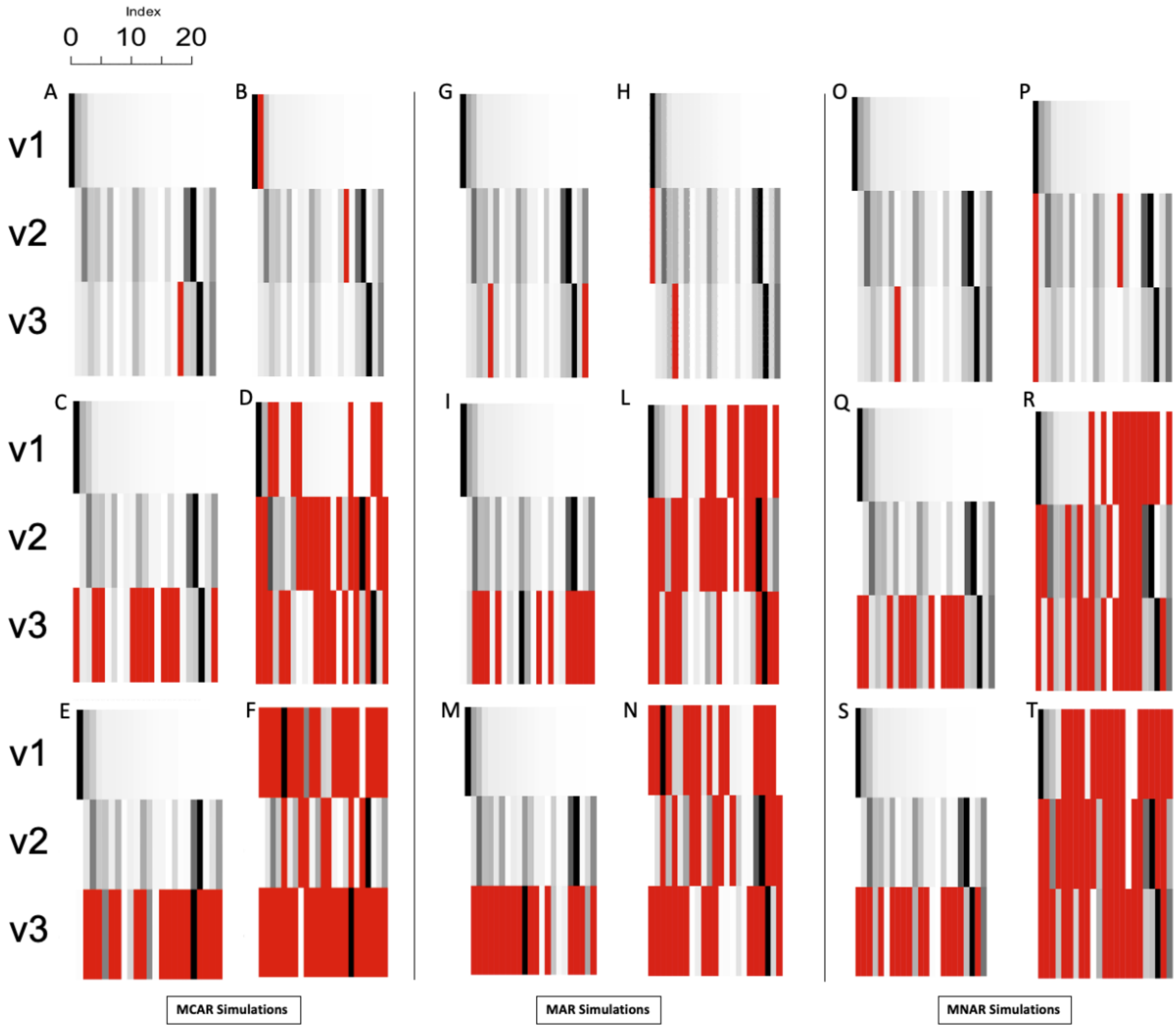

Figure 7. MCAR-MAR-MNAR simulations of synthetic MI distributions. Red pixels indicate synthetic MI distributions in the Au data from Bliss (2004) with darker greys indicating the highest values of tonnage (v1), Au grade (v2) and Ag grade (v3). Data matrices represent a 3 variables model with 23 samples corresponding to the number of deposits (each pixel represents a sample; the data are sorted by tonnage). Six simulations (in each MI, MCAR, MAR and MNAR) were generated to account for the variable proportion of MI (A, B: 10-15%, C, D: 35%, E, F: 55-60% of MI) and consider single (target) variable (A-C-E) and cross-variable cases (B, D, F) with MI distributed in all variables.

A final step of experimentation considered a comparison of imputation methods against listwise deletion to evaluate the difference in Ag grade estimates (see Figures 10, 11 for further details on the results).

*Table 1. Summary of simulation results, top rows (nrow = 30) with less than 10% bias included (data sorted by CR = coverage rate). Total number of experiments ne = 108 with Sim N = 100, corresponding to a total of 10,800 simulations (ne*Sim N). Simulations involved MI over a single (target) variable [001] or equally distributed on three variables [111]. Sampling rate considered resampling with subpopulations with n = 50, 100 samples per simulation. Each experiment considered also variable percentages (p) of MI [15, 35, 55 %] (RB = raw bias, PB = percent bias, CR = coverage).*

nvar MI	n	RB	PB	CR	AW	RMSE	p (%)	MI Type	Sim N	Imputation Method
[111]	100	-0.003	0.97	0.97	1.00	0.23	15	MAR	100	pmm
[001]	50	0.020	7.91	0.97	1.14	0.28	15	MAR	100	norm.nob
[001]	100	0.008	3.05	0.96	0.80	0.19	15	MAR	100	norm.nob
[111]	100	-0.009	3.52	0.96	0.93	0.24	15	MAR	100	norm.nob
[001]	100	0.012	4.67	0.96	0.79	0.19	15	MAR	100	pmm
[001]	100	0.014	5.27	0.96	0.92	0.24	35	MCAR	100	norm.nob
[111]	100	0.021	8.32	0.96	0.95	0.24	15	MCAR	100	pmm
[111]	50	0.005	1.81	0.95	1.75	0.39	35	MCAR	100	norm.nob
[111]	100	0.018	6.92	0.95	0.91	0.24	15	MCAR	100	norm.nob
[001]	100	0.017	6.42	0.94	0.80	0.21	15	MCAR	100	norm.nob
[001]	50	0.023	8.76	0.94	0.97	0.28	15	MAR	100	norm.predict
[111]	50	0.023	9.07	0.94	1.40	0.32	15	MAR	100	norm.nob
[111]	100	-0.001	0.41	0.93	1.27	0.31	35	MCAR	100	norm.nob
[001]	100	0.021	8.04	0.93	0.81	0.21	15	MCAR	100	pmm
[001]	100	0.020	7.62	0.92	0.68	0.21	15	MCAR	100	norm.predict
[001]	50	0.023	8.75	0.92	1.16	0.30	15	MCAR	100	norm.nob
[001]	100	0.009	3.44	0.91	0.67	0.19	15	MAR	100	norm.predict
[111]	50	-0.010	3.91	0.91	1.33	0.32	15	MCAR	100	norm.nob
[001]	100	0.013	5.01	0.91	0.89	0.23	35	MAR	100	norm.nob
[001]	50	0.003	1.34	0.87	1.25	0.41	55	MNAR	100	norm.nob
[001]	100	0.013	4.95	0.87	1.01	0.30	55	MCAR	100	norm.nob
[001]	50	0.017	6.75	0.87	0.96	0.30	15	MCAR	100	norm.predict
[111]	100	-0.001	0.45	0.80	0.76	0.27	15	MNAR	100	norm.predict
[001]	100	0.018	7.03	0.80	0.58	0.23	35	MAR	100	norm.predict
[111]	50	-0.025	9.57	0.80	1.02	0.38	15	MAR	100	norm.predict
[001]	100	0.022	8.41	0.77	0.59	0.24	35	MCAR	100	norm.predict
[001]	100	0.019	7.34	0.59	0.48	0.30	55	MCAR	100	norm.predict
[111]	50	0.003	1.07	0.37	inf	0.93	35	MNAR	100	norm.predict
[111]	100	0.017	6.39	0.28	inf	1.28	55	MAR	100	norm.predict
[111]	100	0.001	0.47	0.26	inf	1.51	55	MCAR	100	norm.predict

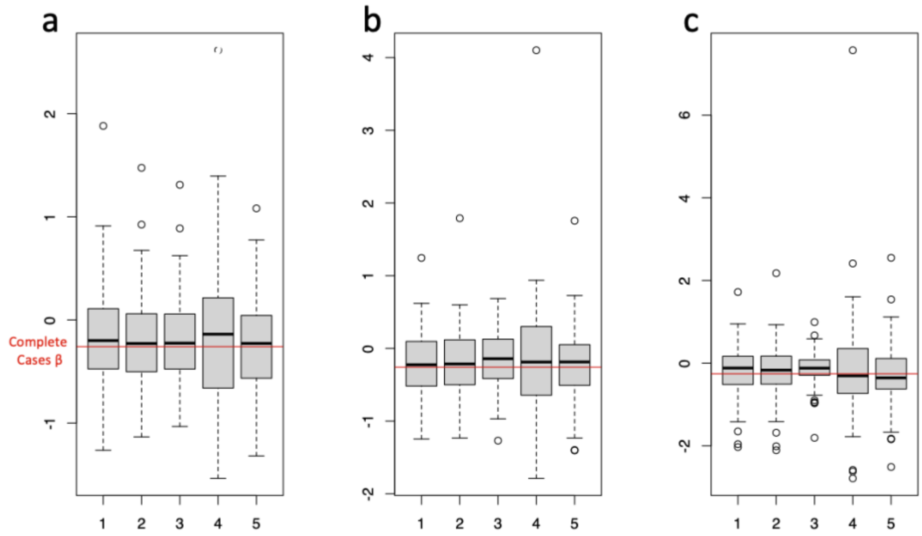

Figure 8. Residuals plots of 500 simulations using MNAR synthetic data (low variable values replaced with MI at variable percentage (a) MI = 15%, (b) MI = 35% and (c) MI = 55% respectively. Boxplots of model performance based on the β regression coefficient estimate derived from synthetic MNAR distributions (100 simulations of MI draws at variable MI percentage and with a sampling rate n = 25). A single variable v3 (Ag_{pct}) was amputed to produce MI-MNAR distributions and subsequently replaced with values obtained from 5 distinct imputation methodologies (1-5) respectively representing 1 = norm.nob, 2 = norm.predict, 3 = pmm, 4 = ml, 5 = norm.mnar.

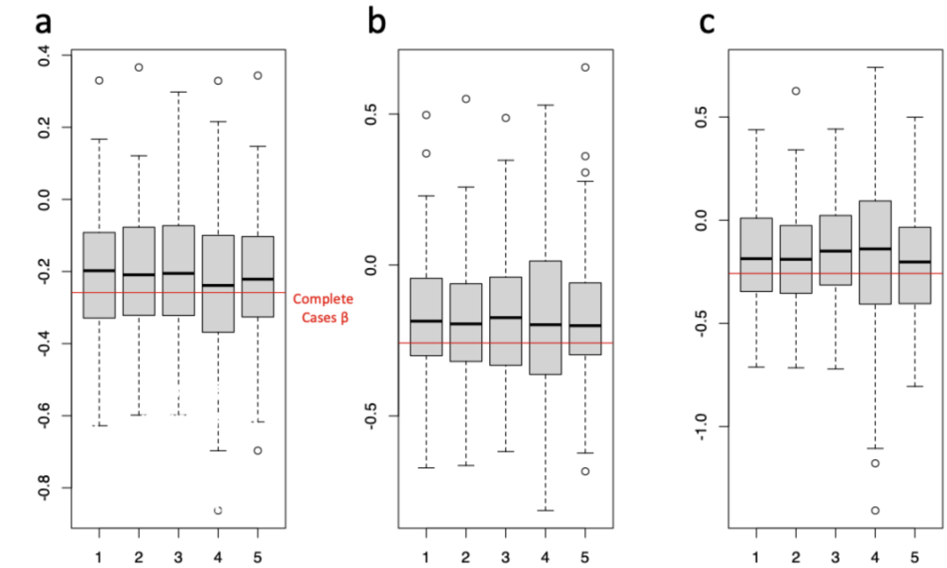

Figure 9. Residuals plots of 500 simulations using MNAR synthetic data (low variable values replaced with MI at variable percentage (a) MI = 15%, (b) MI = 35% and (c) MI = 55% respectively. Boxplots of model performance based on the β regression coefficient estimate derived from synthetic MNAR distributions (100 simulations of MI draws at variable MI percentage and with a sampling rate n = 100). A single variable v3 (Ag_{pct}) was amputed to produce MI-MNAR distributions and subsequently replaced with values obtained from 5 distinct imputation methodologies (1-5) respectively representing 1 = norm.nob, 2 = norm.predict, 3 = pmm, 4 = ml, 5 = norm.mnar.

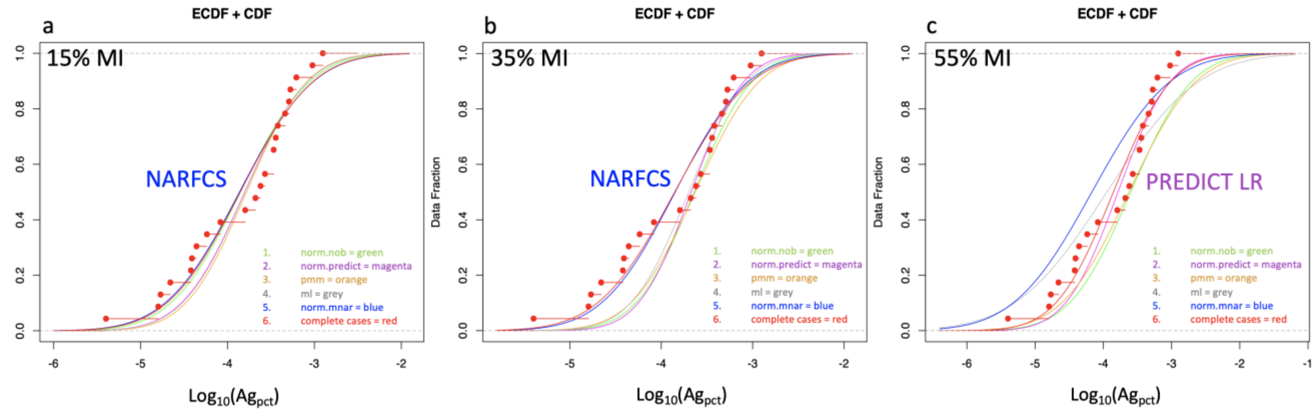

Figure 10. Comparative logistic, cumulative distribution functions for different percentages of MI [(a) 15%, (b) 35% and (c) 55% respectively]. Six cumulative distribution curves are plotted in logarithm scale to allow the comparison of grade prediction models resulting from different imputation strategies (1 to 5) and compared to the complete cases fit (original complete data, regression model 6), the empirical cumulative distribution of complete cases (ECDF) is also reported (red dots). Selected "best" models NARFCS (not at random fully conditional specification, MNAR) and for the latter at 55% MI norm.predict (standard linear regression, PREDICT LR).

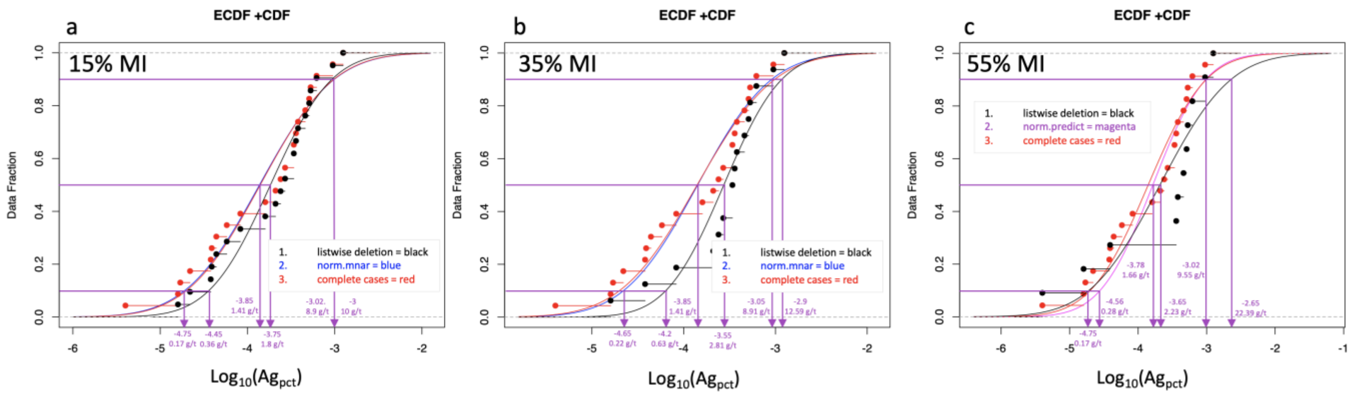

Figure 11. Comparative logistic, cumulative distribution functions for different percentages of MI [(a) 15%, (b) 35%, and (c) 55% respectively]. Three curves are plotted in logarithm scale to allow the comparison of grade prediction models resulting from: (1) listwise deletion of rows containing missing Ag data, (2) the optimal imputation obtained from earlier comparisons, and (3) the complete cases fit (original complete data). Relative empirical distributions (ECDF) are also presented for comparison.

4 DISCUSSION AND CONCLUSIONS

4.1 Factors that should be considered for model selection and their implications

The first aspect that emerges from a closer examination of the results of MI and imputation simulations is that sensitivity analysis is a necessary step when implementing imputation for data augmentation purpose and if solution optimization is a need for the analysis in question. In most cases, sensitivity tests will be part of the imputation approach as the results are heavily dependent on the input data considered, thus they may well change if we consider for instance different ore deposit styles that inherently have unique statistical intercorrelations and MI distributions.

In particular, the assessment of the “best” algorithms for imputing grade and tonnage data in UMRA would be best treated by considering the number of deposits (equating to the sampling rate in the simulations and showing broadly its impact on imputation model performance). For instance, if Figures 8 and 9 (with respectively biases for sampling rates of 25 and 100 synthetic deposits) are compared it appears clear how a larger simulated, sampling population reduces the performance differences across different imputation algorithms. This result suggests that the choice of the methodology will be most relevant when small datasets (small number of deposits) are considered.

In addition, model selection appears to be dependent and requires a close evaluation of the proportion of MI present in the data. This is exemplified by a variation in model performances observed if comparing in both Figures 8, 9 the distances of box plot medians from the complete case fit median value, which will vary dramatically with increasing proportions of MI (cf. cases a, b, c). In this context, it is assumed that lower MI percentages would reflect more reasonable model performances, since the algorithms work best in these scenarios where the statistical correlations are partly preserved in the data.

Another important aspect is to gain an understanding of MI distributions with respect to internal (MNAR) and external variable dependency (MAR). The ideal workflow will then examine the existing data correlations, and the abundance and cross-variable distribution of missing information with respect to internal and external associations. In the Chugach deposit data, we observed weak correlations between Ag and Au data (Pearson correlation = 0.47) with missing information being limited to the Ag variable and having MI in most cases occurring in lower tonnages, indicating a possible link of MI to project maturity or deposit size, suggesting at a minimum a MAR type MI distribution. Further to this, assumptions discussed in Figure 6 which would indicate underestimation bias due to sampling concentration in domains that are favorable to Au endowment led to consider the distribution of MI in Ag grade data likely to be best represented in the MNAR type incompleteness. These considerations are valuable in the selection of imputation models accounting for the specific type of MNAR distributions and led to comparisons carried out using the mnar.norm extension illustrated in Figures 10, 11 where comparative analysis clearly shows improvements in the fit when considering low (up to 15%) to medium (up to 35%) percentages of missing information (Moreno-Betancur and Chavance 2016; Tompsett et al. 2018). In this case, these MNAR specific models demonstrated to be better than multiple imputation pmm and norm2 maximum likelihood, given also that only a single variable is affected by missingness, which tends to favor stochastic imputation (norm.nob, mnar.norm implementations). A close examination of logistic functions representative of model fits compared to the empirical distribution of complete cases (original data before amputation) helps further with the evaluation of model performance and illustrates the tendency of most of the algorithms to lead to overestimation of Ag grades, especially at the lower end of the cumulative distributions, with the only exception represented by the fit considering a correction due to the MNAR type MI. A similar overestimation is also observed if we apply the listwise deletion methodology as illustrated in Figure 11 with an overestimation that is even more pronounced than most of the considered imputation methods (except mnar.norm) and would lead to grade estimates that could be overoptimistic, and likely affecting the UMRA models.

4.2 Conclusive remarks

In summary sensitivity analysis should be run on a specific dataset to evaluate several model fits rather than implementing imputation blindly with a single pass solution, given the expected heterogeneity of grade and tonnage data. With respect to the methodology of choice, the results obtained agree with earlier research indicating that even in these relatively simple mineral resource dataset, stochastic regression imputation and predictive mean matching offer less biased solutions (with pmm being slightly superior because of narrower confidence bounds and much lower raw biases in cases with small sampling rate and with elevated MI, which is attributed to the reduction in variable correlations, with larger MI percentages). In contrast, standard linear regression offers the worst performances with none of the results returning RB values under the 5% threshold in MAR and MCAR analyses (cf. Table 1). If we consider model selection in the context of the Chugach Au-Ag data, it is safe to consider appropriate the use of the mnar.norm algorithms given to relatively moderate amount of MI and its inferred MNAR-type. Despite this appearing to be the optimal selection, these results should not be translated to other resource data of similar nature (for instance applying the same algorithm to another epithermal Au-Ag district dataset). It will always be prudent to complete a series of simulations and compare the results as in some circumstances more than one algorithm could be a viable solution to the imputation problem or as discussed changes in the correlation structure and its association to missing information patterns will condition the results of a specific imputation methodology.

In the more general context of evaluating if imputation could effectively improve the UMRA methodology experimental work suggests that in specific cases (such as MI distributions that conform to the MNAR type) the use of imputation could preserve useful information in the original data records, mitigate the downside effects of listwise deletion method by elimination of bias induced by rows removal, thus providing more accurate estimation of grade and tonnage models.

5 ACKNOWLEDGEMENTS

The authors are grateful to Kalevi Rasilainen for comments and discussion that helped improving this work.

6 REFERENCES

Bliss, J. D. 2004. Grade and tonnage model of Chugach-type low-sulfide Au-quartz veins. In James D. Bliss (Ed.), Developments in Mineral Deposit Modeling: U.S. Geological Survey Bulletin 44–46 p.

Cheema, J. R. 2014. A Review of Missing Data Handling Methods in Education Research. Review of Educational Research, *84*(4), 487–508. https://doi.org/10.3102/0034654314532697

da Silva, C. Z., & Costa, J. F. C. L. 2019. The treatment of missing 'not at random' geological data for ore grade modelling. Applied Earth Science: Transactions of the Institute of Mining and Metallurgy, *128*(1), 15–26. https://doi.org/10.1080/25726838.2018.1547508

Enders, C. K. 2010. Applied missing data analysis. New York, NY, US: Guilford Press.

Ferguson, D., & Winkler, W. 2000. Glossary of Terms on Statistical Data Editing. In Conference of European Statistics Methodological Material, Geneva: United Nations. 18 p.

Geekiyanage, S. C. H., Tunkiel, A., & Sui, D. 2020. Drilling data quality improvement and information extraction with case studies. Journal of Petroleum Exploration and Production 2020 11:2, *11*(2), 819–837. https://doi.org/10.1007/S13202-020-01024-X

Kepezhinskas, P., Berdnikov, N., Kepezhinskas, N., & Konovalova, N. 2022. Adakites, High-Nb Basalts and Copper–Gold Deposits in Magmatic Arcs and Collisional Orogens: An Overview. Geosciences, *12*(1), 1–60. https://doi.org/10.3390/geosciences12010029

Madani, N., & Bazarbekov, T. 2021. Enhanced conditional Co-Gibbs sampling algorithm for data imputation. Computers and Geosciences, *148*(April 2020), 104655. https://doi.org/10.1016/j.cageo.2020.104655

Moreno-Betancur, M., & Chavance, M. 2016. Sensitivity analysis of incomplete longitudinal data departing from the missing at random assumption: Methodology and application in a clinical trial with drop-outs. Statistical Methods in Medical Research, *25*(4), 1471–1489. https://doi.org/10.1177/0962280213490014

Morris, T. P., White, I. R., & Royston, P. 2014. Tuning multiple imputation by predictive mean matching and local residual draws. BMC Medical Research Methodology, *14*(1), 1–13. https://doi.org/10.1186/1471-2288-14-75

Rubin, D. B. 1976. Inference and Missing Data. Biometrika, *63*(3), 581. https://doi.org/10.2307/2335739

Schafer, J. 1999. Multiple imputation: A primer. Statistical Methods in Medical Research, *8*(1), 3–15. https://doi.org/10.1191/096228099671525676

Schafer, J. 2021. Package “norm2” Title Analysis of Incomplete Multivariate Data under a Normal Model. https://cran.r-project.org/web/packages/norm2/index.html. Accessed 23 November 2022

Singer, D. A. 1993. Basic concepts in three-part quantitative assessments of undiscovered mineral resources. Nonrenewable Resources, (2), 69–81.

Singer, D., & Menzie, W. D. 2010. Quantitative Mineral Resource Assessments: An Integrated Approach. New York: Oxford University Press. https://doi.org/10.1093/oso/9780195399592.001.0001

Takahashi, M. 2017. Statistical inference in missing data by MCMC and non-MCMC multiple imputation algorithms: Assessing the effects of between-imputation iterations. Data Science Journal, *16*, 1–17. https://doi.org/10.5334/dsj-2017-037

Tompsett, D. M., Leacy, F., Moreno-Betancur, M., Heron, J., & White, I. R. 2018. On the use of the not-at-random fully conditional specification (NARFCS) procedure in practice. Statistics in Medicine, *37*(15), 2338–2353. https://doi.org/10.1002/sim.7643

van Buuren, S. 2018. Flexible imputation of missing data Second Ed. Boca Raton, FL: CRC Press, Taylor & Francis Group.

van Buuren, S., & Groothuis-Oudshoorn, K. 2011. mice: Multivariate imputation by chained equations in R. Journal of Statistical Software, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Wilks, S. 1932. Moments and Distributions of Estimates of Population Parameters. Annals of Mathematical Statistics, (3), 163–195. https://doi.org/10.1214/aoms/1177732885